The Wage Gap between Metropolitan and Non-metropolitan Areas

Bonggeun Kim[*]

September 30, 2003[**]

Abstract: In the literature on measured wage inequality, only one recent study, by Glaeser and Mare′ (2001), has focused on the enormous wage gap between urban and non-urban workers in the United States. In the present paper, I replicate and extend Glaeser and Mare′'s original empirical work, and I present a new interpretation of the evidence based on my re-estimation. Contrary to Glaeser and Mare′'s theory that urban employment induces more rapid skill acquisition, I find that wage growth is no greater for urban workers than for non-urban workers. I show that both the original and extended empirical patterns can be fully explained by a simple spatial equilibrium model that incorporates two highly plausible phenomena: (1) a compensating wage differential for the higher cost of living in cities and (2) a dynamic tendency for more able workers to gravitate to cities once they discover that they belong in the "big leagues."

Key words: urban/non-urban wage gap, real wage gap, dynamic ability sorting, and market learning

JEL Classification: J31

---

* Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, Tel: (00643) 364-2823, email: bonggeun.kim@canterbury.ac.nz, Fax: (00643)364-2635
** I am grateful to Gary Solon, George Johnson, Gordon Hanson, Rebecca Blank, and participants at the labor seminar at the University of Michigan for their helpful comments.

The Wage Gap between Metropolitan and Non-metropolitan Areas

I. Introduction

In a cross-sectional comparison of men with the same years of education and work experience, the wages received by workers in metropolitan areas are typically about 30 percent higher than those received by non-metropolitan workers. When large metropolitan areas with a population of at least 1 million are compared to non-metropolitan areas, the wage gap is even larger, more than 40 percent in some data sets. This gap is enormous, larger than many other heavily researched wage gaps, such as the union/non-union wage gap; however, the urban/non-urban wage gap has been almost completely overlooked in the literature on measured wage inequality in the United States.[1] There is a well-developed literature on wage differences across particular cities,[2] but only one recent study, by Glaeser and Mare′ (2001), has focused on the wage difference between cities and non-cities.

Glaeser and Mare′′s empirical analysis establishes three empirical regularities: (1) the cross-sectional wage gap between cities and non-cities is large; (2) it is larger for more experienced workers than for less experienced workers; and (3) longitudinal estimates of the urban/non-urban wage gap that control for worker fixed effects are much

[1] For less developed countries, see Harris and Todaro (1970). In their seminal paper, they speculate that rural-to-urban migration depends on the urban/rural difference in expected earnings. The urban/rural wage gap is explained by the urban minimum wage, which is institutionally determined to be higher than that of the free market. This wage gap is accompanied by a higher unemployment rate in the urban sector. In the United States, however, there was hardly any difference in unemployment rates between urban and non-urban areas in 1990.

smaller than the cross-sectional estimates. Glaeser and Mare′ interpret these empirical patterns as suggesting that urban employment is conducive to especially rapid skill acquisition. Surprisingly, however, they never subject that hypothesis to a test of its most obvious empirical implication – that urban workers exhibit steeper wage growth than non-urban workers.

My own empirical analysis uses data from the decennial census, the Current Population Survey, and the Panel Study of Income Dynamics (PSID). After replicating the three regularities emphasized by Glaeser and Mare′, I proceed to use the longitudinal data from the PSID to check their theory's implication that wage growth is more rapid for urban workers than for non-urban workers. The data show no support for that hypothesis.

I therefore consider whether the entire set of empirical regularities might be explained by an alternative theory. I show that all the empirical patterns can indeed be explained by a simple spatial equilibrium model that incorporates two highly plausible phenomena: (1) a compensating wage differential for the higher cost of living in cities and (2) a dynamic tendency for more able workers to gravitate to cities once they discover that they belong in the "big leagues."

The next section of this paper provides a brief review of the relevant theoretical and empirical literature. In Section III, I describe my data and present my cross-sectional and longitudinal evidence. In Section IV, I develop my theoretical model and demonstrate that it explains all the empirical patterns established in Section III. Section V concludes with a summary of the main findings.

---

[2] See Roback (1982), Johnson (1983), and Rauch (1993).

II. Literature Survey

To understand the enormous urban/non-urban wage gap as an economic equilibrium, we need to answer two fundamental questions: (1) what factors prevent all the workers from flocking to the urban areas, and (2) why do urban employers remain in urban areas despite higher wages? As Glaeser and Mare′ note, some of the possible answers to these questions have been foreshadowed in the related literature on wages differences across different cities. As I discuss the answers suggested in the previous literature, I also will note their empirical implications, which I will test in Section III.

Starting with question (1), non-urban workers may be deterred from migrating to cities, despite the higher urban wages, because of urban disamenities, especially the higher urban cost of living. Roback (1982) and Rauch (1993) suggest that the utility level of workers is equal across cities with different wages because the intercity wage gaps merely compensate for the intercity cost-of-living differences. Similarly, if the urban wage premium is merely a compensating difference for a higher cost of living, it does not induce all the non-urban workers to want to migrate to the cities. While this answer to question (1) is theoretically coherent, Page and Solon (2003) emphasize that its empirical plausibility depends on whether the urban/non-urban cost-of-living difference is large enough to account for the large urban/non-urban wage gap. In Section III, I will explore whether it is.

Another answer to question (1), suggested long ago by Johnson (1953), is that the urban workers typically have higher ability than the non-urban workers. To the extent that the high urban wages are a return to the high ability of the urban workers, less able non-urban workers have no incentive to migrate to the cities because, even if they did, they would not have the ability to command the higher wages. This is a theoretically coherent explanation provided there is some reason for the high-ability workers to have a comparative advantage for urban employment. One possibility is that the density of urban markets generates a larger clientele for the top talent. It was an equilibrium for the best basketball player to work in Chicago, rather than in the hinterland, because he could reach a larger audience there. For similar reasons, it is unsurprising that many of the best ballet dancers and lawyers are drawn to New York City.[3] Again there is a straightforward empirical implication. If much of the observed urban/non-urban wage gap is a return to ability (or other worker-specific characteristics like motivation), then longitudinal estimates of the wage gap that control for worker "fixed effects" should be smaller than the cross-sectional estimates. Again I will test this implication in Section III.

If the urban/non-urban wage gap is solely a consequence of ability differences between urban and non-urban workers, then question (2) is not a puzzle after all. Urban employers can afford to pay higher wages if they are compensated by the greater productivity of their high-ability workers. On the other hand, if the urban wage premium is not completely explained by ability differences (e.g., if a substantial urban/non-urban wage gap is estimated even after controlling for worker fixed effects), question (2) still requires an answer. In that case, urban employers could thrive in the face of higher labor costs if that cost disadvantage is offset by various cost or productivity advantages. One of these is lower transportation costs (Krugman, 1991). Another possible economy of agglomeration, emphasized by Lucas (1988), is the productivity advantage from greater knowledge spillovers in urban areas.

[3] Rosen (1981) makes a similar argument for why the most able workers should rise to the top of organizational hierarchies.

Glaeser and Mare′ espouse a variation on Lucas's theme. If the knowledge spillovers operate on individual workers rather than firms, then urban workers, even if they are not initially more able, might gradually accumulate more human capital than non-urban workers. One empirical implication, corroborated by Glaeser and Mare′, is that the cross-sectional wage gap between cities and non-cities is greater for more experienced workers. Another, which Glaeser and Mare′ do not report checking, is that longitudinal wage growth is greater for urban workers than for non-urban workers. I check both empirical implications in the next section.

III. Empirical Evidence

3.1. Data

I use data from the 1990 census 1 percent Public Use Microdata Samples (PUMS), the 1991 March Current Population Survey (CPS), and the Panel Study of Income Dynamics (PSID). I use the census and the CPS for cross-sectional estimation only, and I use the PSID for both cross-sectional and longitudinal estimation. The census data provide the largest sample, and the CPS is used here to check the robustness of the results from the other data sets. The PSID is a longitudinal survey conducted by the University of Michigan Survey Research Center. I exclude the Survey of Economic Opportunity sample to eliminate oversampling of poor families. In all data sets, I restrict my samples to male heads of household.[4] I use only positive earners between the ages of 18 and 65.

[4] I also have investigated female heads of household with the same restrictions, and the wage patterns of female heads across areas are extremely similar to those of male heads. For consistency with Glaeser and Mare′'s analysis, I restrict my census and CPS samples to workers who usually worked at least 35 hours per week over the previous year.

My wage variable is average hourly earnings, constructed as annual labor earnings divided by annual hours of work. My main analysis of the PSID uses a balanced panel of seven years of data from 1985 to 1991 for the same head of household. Reported wage information is based on the previous calendar year, so earnings observations come from the interviews for 1986 to 1992. Thus, for 1991, wage information ascertained from the 1992 interview is the basis for the dependent variable, and personal characteristics and regional information from the 1991 interview are the basis for the explanatory variables. My main longitudinal analysis is limited to the years 1985-1991 because I use the new metropolitan status variable that first became available in 1985. The PSID metropolitan status variable used in Glaeser and Mare′ is based on the 1960 census, which induces substantial measurement error due to outdated metropolitan status. The new metropolitan variable is based on the 1980 census.[5]

The concept of a Metropolitan Area (MA) is a large population nucleus together with adjacent communities that have a high degree of economic and social integration with that nucleus. Each MA must contain either a place with a minimum population of 50,000 or a Census Bureau-defined urbanized area and a total MA population of at least 100,000. The MA is subdivided into "central city" and "suburban areas." The areas located outside the metropolitan areas are referred to as non-metropolitan areas. The standards provide a flexible structure of metropolitan definitions that classify a metropolitan area either as a metropolitan statistical area (MSA) or as a consolidated metropolitan statistical area (CMSA) that is divided into primary metropolitan statistical areas (PMSA's). In the present paper I use MSA/PMSA as metropolitan areas in the

[5] For comparison purposes, I also try using the same first 18 waves of the PSID that Glaeser and Mare′ did with their metropolitan status indicator.

census and the CPS, and I use MA in the PSID. The census and the CPS metropolitan areas are based on the city-level data, but the metropolitan areas in the PSID are based on the county-level data. To examine wage differences between metropolitan areas of different sizes, I will divide metropolitan areas into large metropolitan areas of 1 million or more and smaller metropolitan areas. Furthermore, I will investigate wage patterns between central cities and suburban areas in large metropolitan areas.

Aside from the metropolitan/non-metropolitan indicators, the explanatory variables will include region dummies, a race dummy, a polynomial in potential experience (age-years of schooling-6), and education dummies as defined in Table 1. Table 1 lists the code number of each PSID variable and the name of each census variable,[6] and Table 2 shows the variables' sample means and standard deviations.

The sample proportions living in metropolitan areas are very similar across data sets. The proportion of the census sample recorded as living in a metropolitan area is 77.6 percent. The log point urban/non-urban wage gap is 0.30. In the PSID, 71.3 percent of the sample lives in a metropolitan area, and the wage gap is 0.41. The CPS shows a similar proportion of metropolitan residents and log point urban/non-urban wage gap as in the census, which is not consistent with Glaeser and Mareʹ.[7]

---

[6] The variable names in the CPS are very similar to those for the census, so I list only the census names.

[7] The census data also include information regarding "place of work," which enables study of the wage patterns of commuters between areas. In the census, 79.4 percent work in metropolitan areas, which is slightly higher than the percentage that resides in metropolitan areas. This discrepancy implies the existence of workers commuting from non-metropolitan areas to metropolitan areas. The existence of commuting workers, who comprise only a few percent of each sample, is not of much consequence for most of my subsequent cross-sectional and longitudinal estimation. For example, in the census, if I classify metropolitan workers according to place of work instead of residence, the estimate of the metropolitan coefficient in my wage equation increases from 0.222 to 0.233.

3.2. Cross-sectional estimation

In this section, I use all three data sets to study the cross-sectional urban/non-urban wage gap. In the simplest analyses, I apply ordinary least squares to the linear regression model

$$\ln w_i = \beta' X_i + \delta D_i + \varepsilon_i \qquad (3\text{-}1)$$

where $\ln w_i$ is worker i's log hourly wage rate, X_i is a vector of individual characteristics and region dummies, and D_i is a metropolitan dummy variable equal to 1 if the individual lives in a metropolitan area. My census and CPS data sets are cross-sections for a single year. In my cross-sectional analysis of the PSID, the dependent variable is

$$\overline{\ln w_i} = \sum_{t=1985}^{1991} \ln w_{it} / 7 , \qquad (3\text{-}2)$$

the seven-year average of the individual's log wage, and the explanatory variables are similarly averaged over time. Unlike Glaeser and Mare′'s supposedly cross-sectional analysis, which mixes both the "between" and "within" variation in pooled annual data from the PSID, my cross-sectional analysis uses only the "between" variation across individuals and therefore is more comparable to the census and CPS analyses.

The first column of Table 3 shows the coefficient estimates from the census data. The estimated coefficient of the metropolitan dummy, $\hat{\delta} = 0.222$, implies a metropolitan/non-metropolitan wage gap of 25 percent (calculated as $e^{0.222} - 1$). The corresponding estimate from the CPS in the third column implies a wage gap of 26 percent, and the PSID-based estimate in the fifth column implies a gap of 36 percent. Thus, the estimated urban/non-urban wage gap ranges from 25 percent to 36 percent.

The specifications in the second, fourth, and sixth columns of Table 3 allow separate coefficients for large metropolitan areas (containing at least a million residents) and small ones. In all three data sets, the coefficient estimates are larger for large metropolitan areas than for small ones. In the census, for example, the coefficient estimate is 0.27 for large metropolitan areas and 0.16 for small ones.

My estimates from the census and CPS are consistent with Glaeser and Mare's census-based estimates. At first, however, my PSID-based estimates seem larger than theirs. As mentioned above, Glaeser and Mare′'s estimates come from applying OLS to pooled longitudinal data. This estimator is a weighted average of the "between" estimator that I use in my cross-sectional analysis and the "within" estimator that controls for worker-specific fixed effects. As I will show in my longitudinal analysis in the next sub-section, the within or fixed-effects estimate turns out to be much smaller than the between estimate, mainly because the within estimate controls for unobserved ability differences between urban and non-urban workers. By mixing the within variation with the between variation, Glaeser and Mare′ cause their supposedly cross-sectional estimates from the PSID to come out smaller than mine.

Table 4 reports a series of regression analyses that facilitate comparison between my PSID results and Glaeser and Mare′'s. Using both my 1985-1991 panel and an unbalanced panel from Glaeser and Mare′'s 1968-1985 period (along with their metropolitan indicator), I estimate equation (3-1) both with the between estimator and with OLS applied to the pooled individual observations from all years. As shown in the last column, switching from my 1985-1991 period to Glaeser and Mare′'s 1968-1985 period only slightly reduces the between estimate of the metropolitan coefficient, from

0.308 to 0.294. The main reason for their smaller estimates is that their OLS estimator mixes the genuinely cross-sectional between variation with the within variation used by the fixed-effects estimator. As shown in the first column of the bottom panel, using the OLS estimator instead of the between estimator with the 1968-1985 data reduces the estimated metropolitan coefficient from 0.294 to 0.242.[8] The second column of Table 4 highlights one other problem with Glaeser and Mare′′s OLS estimation – their standard error estimation falsely assumes that observations of the same worker in different years are independent. The second column reports Huber/White standard error estimates that account for serial correlation among the error terms in different years for the same workers. These corrected standard error estimates are more than twice as large as the uncorrected ones.

As discussed in Section II, one likely reason for the urban/non-urban wage gap is a compensating differential for the higher cost of living in urban areas. Table 5 reports a series of analyses that explore the importance of this factor. In the first, third, and fifth columns (labeled LHS for adjustment of the left-hand-side variable), I reestimate equation (3-1) with a new dependent variable, the log wage deflated by a local cost-of-living index. This variable is constructed as

$$\ln w_{real,i} = \ln w_i - \theta \ln p_i \tag{3-3}$$

where p_i is an index measuring interarea variation in housing prices and θ is the typical expenditure share devoted to housing. In Section IV below, I demonstrate formally that the expression in equation (3-3) is an appropriate way to adjust for variation in housing

[8] In the top part of the table, switching from the between estimator to OLS causes a smaller reduction because, with only a seven-year panel, the OLS estimator gives less weight to the within variation than it does with Glaeser and Mare′′s eighteen-year panel.

costs. In most of my analyses, I set θ at 0.25 because the Bureau of Labor Statistics Consumer Expenditure Survey for 1990 reported that the average household spent 24 percent of its gross income on shelter and utilities.

The reason for using only the housing component in regional cost-of-living variation is that variation in housing prices is the main source of interarea differences in cost of living according to the National Research Council's Panel on Poverty and Family Assistance.[9] In any case, more general cost-of-living statistics for both metropolitan and non-metropolitan areas are not available. I measure p_i with the cost-of-housing (including utilities) index values reported by the National Research Council panel.[10] These values are categorized (relative to 1 for the United States as a whole) by region (census division) and urbanicity.

Comparing the LHS results in Table 5 to the corresponding results in Table 3 shows that, in all three data sets, adjusting the wage variable for differences in cost of living does reduce the estimated metropolitan coefficient by about 0.1. This reduction is statistically significant in all three data sets. Despite the reduction, the metropolitan coefficient estimate remains statistically significant and substantially positive, at about 0.13 in the census, 0.13 in the CPS, and 0.20 in the PSID. These results suggest that compensation for the higher urban cost of living is part, but only part, of the explanation for the urban/non-urban wage gap.

In reality, other consumption goods besides housing, such as labor services and property taxes, might be more expensive in urban areas. Thus, adjusting for only housing

---

[9] Citro and Michael (1995), p. 183.

[10] Citro and Michael (1995), Table 3-6.

price variation may under-adjust for the cost-of-living variation.[11] I have performed two experiments to consider the possibility of under-adjustment. First, I have checked the sensitivity of my results to varying the value of θ. Raising θ within a reasonable range reduces, but does not eliminate, the estimated urban/non-urban wage gap. In the census data, for example, using a θ of 0.33 instead of 0.25 reduces the estimated metropolitan coefficient from 0.129 to 0.099. Second, in the second, fourth, and sixth columns of Table 5, instead of adjusting the left-hand-side variable with an assumed value of θ, I use $\ln p_i$ as an additional right-hand-side variable with an unconstrained coefficient. If workers with higher unobserved ability tend to locate in more expensive areas, this approach will over-adjust for cost of living, but it seems worthwhile to check how the results are affected. This approach does further reduce the estimated metropolitan coefficient, but, in two of the three data sets, the estimated coefficient remains positive and statistically significant.

The census and the CPS sub-divide metropolitan area into central cities, suburban areas, and other metropolitan areas. I also observe central counties of metropolitan areas of 1 million or more in the PSID. The PSID provides a more restricted definition of central cities because these areas are also in the metropolitan areas of 1 million or more. In other words, I observe only central counties and fringe counties in large metropolitan areas. Table 6 reports results from specifications that distinguish these different portions of metropolitan areas. In the PSID, the coefficient estimates are 0.37 for central cities and 0.43 for suburban areas. The CPS and census estimates also are greater for suburban

[11] On the other hand, as Page and Solon (2003) note, the lower transportation costs and superior consumer amenities in metropolitan areas imply that adjusting only for housing costs may *over*-adjust for the cost-of-living difference between metropolitan and non-metropolitan areas.

areas than for central cities. The place-of-work information in the census clearly shows that the commuters have higher wages. As shown in the second column of the table, individuals who work in central cities tend to have higher wages than those who work in suburban areas.

In Table 7, following Glaeser and Mare′, I examine how the urban/non-urban wage gap varies with education and experience. F-tests of the hypothesis of zero coefficients for the interactions of the metropolitan dummy with the education and experience variables reject that hypothesis. In all three data sets, the urban/non-urban wage gap is greater at higher levels of education. In the census, for example, as shown in the table and in Figure 1, the estimated metropolitan coefficient is 0.15 greater at 16 years of education than at 12 years of education.

Like Glaeser and Mare′, I also find that the urban/non-urban wage gap is greater for more experienced workers. Figure 2, for example, shows that in the census data the metropolitan coefficient is estimated to be more than 0.1 greater at ten years of experience than at zero experience.

To summarize, my cross-sectional analyses of the census, CPS, and PSID corroborate Glaeser and Mare′'s finding of a very large wage gap between metropolitan and non-metropolitan areas. I have extended their work by adjusting for cost-of-living differences, and my results suggest that compensation for the higher cost of living in urban areas accounts for only part of the wage gap. I also have replicated Glaeser and Mare′'s finding that the urban/non-urban wage gap is greater for more experienced workers.

3.3. Longitudinal estimation

The cross-sectional estimates in Section 3.2 suggest that urban wages exceed non-urban wages by more than can be explained solely by differences in cost of living. As discussed in Section II, one possible reason that urban workers tend to earn higher real wages is that they typically are more able. If so, one would expect that longitudinal estimates of the wage gap that control for worker "fixed effects" would turn out to be substantially smaller than the cross-sectional estimates. In this sub-section, I use the longitudinal PSID data from 1985-1991 to explore that implication.

First, I control for worker fixed effects by estimating the "within" regression, that is, by applying OLS to the mean-differenced equation

$$\ln w_{it} - \overline{\ln w_i} = \beta'(X_{it} - \overline{X_i}) + \delta(D_{it} - \overline{D_i}) + \varepsilon_{it} - \overline{\varepsilon_i} \qquad (3\text{-}4)$$

where, for any variable Z, $\overline{Z_i} = \sum_{t=1985}^{1991} Z_{it} / 7$. As shown in the first column of Table 8, this produces an estimated metropolitan coefficient of $\hat{\delta} = 0.113$.[12] As conjectured, this estimate is far below the corresponding cross-sectional estimate of 0.308 in Table 3, suggesting that much of the latter estimate is really a return to unobserved ability. Interestingly, the within estimate of 0.113 appears to just compensate for the 0.1 urban/non-urban cost-of-living difference estimated in the previous sub-section. The second column of Table 8 addresses this point directly by re-estimating equation (3-4) with the wage variable deflated for cost of living. The resulting estimate of the metropolitan coefficient is 0.008, almost exactly zero. That is, the combination of

[12] The reported standard error estimates are the conventional type that assumes no serial correlation remains in the error term once the fixed effects have been controlled for. Using Arellano's (1987) covariance matrix estimator that is robust to serial correlation produces modestly larger standard error estimates.

accounting for cost-of-living differences and accounting for workers' unobserved ability with fixed effects almost exactly explains the entire estimated urban/non-urban wage gap.

An alternative explanation for why the within estimate is less than the between estimate is that within estimation magnifies the downward errors-in-variables bias from classification error in the metropolitan dummy. To examine this possibility, I follow Griliches and Hausman's (1986) suggestion to control for fixed effects by applying OLS to the long-differenced, rather than mean-differenced, equation

$$\ln w_{i,1991} - \ln w_{i,1985} = \beta'(X_{i,1991} - X_{i,1985}) + \delta(D_{i,1991} - D_{i,1985}) + \varepsilon_{i,1991} - \varepsilon_{i,1985}. \quad (3\text{-}5)$$

The idea is that measuring change in metropolitan status over a longer time period captures more "signal" relative to "noise" in the measured change and therefore reduces the errors-in-variables bias. As shown in the third column of Table 8, this approach does indeed increase the estimated metropolitan coefficient from 0.113 to 0.142. Still, this estimate remains much smaller than the cross-sectional estimate.

The long-differenced specification in equation (3-5) is readily extended to the more general model

$$\ln w_{i,1991} - \ln w_{i,1985} = \beta'(X_{i,1991} - X_{i,1985}) + \delta_1 D_{i,nm} + \delta_2 D_{i,mn} + \delta_3 D_{i,mm}$$

$$+ \varepsilon_{i,1991} - \varepsilon_{i,1985} \quad (3\text{-}6)$$

where $D_{i,nm}$ is a dummy variable that equals 1 if, between 1985 and 1991, worker i moved from non-metropolitan to metropolitan; $D_{i,mn}$ equals 1 if he moved from metropolitan to non-metropolitan; and $D_{i,mm}$ equals 1 if he lived in a metropolitan area in both 1985 and 1991. The three δ coefficients represent the relative wage growths of the three categories compared to the omitted category of staying in a non-metropolitan area.

Relative to this generalized model, the model in equation (3-5) imposes the restrictions $\delta_3 = 0$ and $\delta_2 = -\delta_1$.

As shown in the last column of Table 8, OLS estimation of equation (3-6) yields $\hat{\delta}_1 = 0.187$, $\hat{\delta}_2 = -0.097$, and $\hat{\delta}_3 = -0.013$. Thus, the wage gap suggested by the wage growth of movers from non-metropolitan to metropolitan areas is somewhat larger than the 0.142 value estimated in the third column, while the gap suggested by the movers from metropolitan to non-metropolitan areas is somewhat smaller. The most striking result, however, is the estimate of δ_3. Glaeser and Mare′ claim that the urban/non-urban wage gap arises because urban employment is conducive to more rapid human capital accumulation. If that were the real story, stayers in metropolitan areas should exhibit more rapid wage growth than stayers in non-metropolitan areas, i.e., δ_3 should be positive. The insignificantly negative estimate of δ_3 indicates no empirical support whatsoever for Glaeser and Mare′'s interpretation.[13]

IV. The Model

The previous section documented a series of empirical regularities. My cross-sectional estimates indicate that the urban/non-urban wage gap is very large, exceeds the amount necessary to compensate for the higher urban cost of living, and is larger among more experienced workers. My longitudinal estimates, which control for worker fixed effects, are smaller than the cross-sectional estimates and are just sufficient to

[13] I obtain similar results after excluding workers who lived in a metropolitan (or non-metropolitan) area in both 1985 and 1991, but had switched back and forth between 1985 and 1991. In a somewhat related analysis, Glaeser and Mare′ investigate the time pattern of wage

compensate for the higher urban cost of living. Contrary to Glaeser and Mare′’s story about more rapid skill acquisition in cities, workers in metropolitan areas exhibit no more wage growth than workers in non-metropolitan areas. In this section, I develop a simple spatial equilibrium model that accords with this entire set of empirical regularities.

In my model, more able workers have a comparative advantage in the urban sector and therefore tend to concentrate in cities. Thus, the cross-sectional wage gap between urban and non-urban areas exceeds the cost-of-living difference because it is partly a return to the higher ability of urban workers. Longitudinal estimates of the wage gap that control for worker fixed effects come out smaller because they control for unobserved ability. My model, like Gibbons and Katz’s (1992) model of wage differences and worker sorting among industries, assumes that workers are not at first fully aware of their ability levels, but learn them over time. The ability sorting of workers between the urban and non-urban sectors therefore is a gradual process, with the ability gap between the two sectors growing as workers gravitate to their sector of comparative advantage. As a consequence, the wage gap between the two sectors is larger for more experienced workers. This gradual sorting process can also be explained by time-consuming search process. Workers and firms could have full information about the workers’ ability types, but the workers’ efficient allocation might be gradually achieved due to search or mobility cost. The gradual sorting by learning is one of theoretical interpretations which are consistent with the data. The remainder of this section presents a simple formalization of this story.

---

changes for movers to and from metropolitan areas. Their results from both the PSID and the NLSY show no clear tendency for greater wage growth in the years following a move to a city.

4.1. Worker preferences

Consider a stationary population of overlapping generations, each of which works for two periods (t=1, 2). In each period, each worker i chooses between working in the metropolitan sector ($D_{it} = 1$) or the non-metropolitan sector ($D_{it} = 0$).

The worker's valuation of working in each sector follows the indirect utility function $V = \ln w - \alpha D$ where w is the worker's opportunity wage in that sector. Thus, the worker would choose the metropolitan sector if and only if his metropolitan log wage exceeds his non-metropolitan log wage by at least α, the metropolitan wage premium needed to compensate for the higher cost of living (or other metropolitan disamenities net of the amenities).

This indirect utility function is consistent with a utility maximization in which the worker maximizes the Cobb-Douglas utility function $(1-\theta)\ln X + \theta \ln H$ subject to the budget constraint $w = X + pH$, where X is the worker's consumption of a traded good (which is treated as the numeraire), H is his consumption of a non-traded good (H for housing), and p is the sector's price of the non-traded good. It is straightforward to show that the resulting indirect utility function is $V = constant + \ln w - \theta \ln p$. Apart from the constant, this is the same as the assumed indirect utility function with $\alpha = \theta \ln(p_m / p_n)$ where p_m / p_n is the housing price ratio between the metropolitan and non-metropolitan sectors. Note that this result accords with the method of price deflation used above in Section III.

4.2. Productivity by sector and ability

Although worker preferences are homogeneous, worker abilities are not. Workers are of two types, high-ability and low-ability, and they contribute differently to each sector's production of the traded good X .

In each sector j ($j = m, n$), many competitive firms produce X with a constant-returns-to-scale technology, which aggregates to the sectoral production function

$$X_j = T_j F(K_j, \eta_{Hj} N_{Hj} + \eta_{Lj} N_{Lj}) \qquad (4\text{-}1)$$

where T_j is a sector-specific productivity factor, K_j is the capital input (including land), N_{Hj} and N_{Lj} are the sector's numbers of high- and low-ability workers, and η_{Hj} and η_{Lj} represent the relative productivities of the two ability types in sector j . If cities enjoy some of the productivity advantages discussed in Section II, $T_m > T_n$. The greater productivity of high-ability workers in both sectors is expressed by $\eta_{Hj} > \eta_{Lj}$, and the comparative advantage of high-ability workers in urban employment is expressed by $(\eta_{Hm} / \eta_{Lm}) > (\eta_{Hn} / \eta_{Ln})$.

At the time young workers choose their sector in $t = 1$, they (and the employers) are ignorant of their ability types. All that is (commonly) known at that point is that a proportion π of the population has high ability. Immediately upon the initiation of production in period 1, each worker's ability becomes publicly known. This knowledge informs the employers' decisions about wages in both periods, and it informs the workers' sectoral choices for the second period.

4.3. Labor market equilibrium

With competitive markets, the employers equate wages to value of marginal product. Consequently, $w_{Hj} = T_j F_{2j} \eta_{Hj}$, $w_{Lj} = T_j F_{2j} \eta_{Lj}$, and thus $(w_{Hj} / w_{Lj}) = (\eta_{Hj} / \eta_{Lj})$. Given the previous sub-section's assumptions about relative productivities, it follows that

$$\frac{w_{Hm}}{w_{Lm}} > \frac{w_{Hn}}{w_{Ln}} > 1 \tag{4-2}$$

or, equivalently, that

$$\ln w_{Hm} - \ln w_{Lm} > \ln w_{Hn} - \ln w_{Ln} > 0 . \tag{4-3}$$

Now consider an equilibrium in which young workers select into both the metropolitan and non-metropolitan sectors. Given that the workers are initially ignorant about their ability levels, this requires that they be indifferent between the two sectors and choose between them randomly. Thus, if the workers are risk-neutral and maximize the expected value of the utility function in Section 4.1,

$$\alpha = E(\ln w_m) - E(\ln w_n) = \pi \ln w_{Hm} + (1-\pi)\ln w_{Lm} - [\pi \ln w_{Hn} + (1-\pi)\ln w_{Ln}]$$

$$= \pi(\ln w_{Hm} - \ln w_{Hn}) + (1-\pi)(\ln w_{Lm} - \ln w_{Ln}) . \tag{4-4}$$

Since α equals a weighted average of the metropolitan/non-metropolitan log wage gaps of the two ability types, and since (4-3) implies that the high-ability gap exceeds the low-ability gap, it follows that

$$\ln w_{Hm} - \ln w_{Hn} > \alpha \tag{4-5}$$

and

$$\ln w_{Lm} - \ln w_{Ln} < \alpha . \tag{4-6}$$

Note that inequalities (4-5) and (4-6) determine the sectoral sorting of workers in period 2, when the workers are aware of their ability levels and gravitate to their sector of comparative advantage. High-ability workers sort into the metropolitan sector because their metropolitan wage premium exceeds the cost-of-living differential α, and low-ability workers choose the non-metropolitan sector because their metropolitan wage premium falls short of α.

We are now in a position to collect the full set of implications for cross-sectional and longitudinal wage patterns. Starting with cross-sectional implications, equation (4-4) tells us that, for young workers, the average log wage gap between the metropolitan and non-metropolitan sectors equals the cost-of-living differential α. For older workers, who have sorted by ability, inequalities (4-5) and (4-2) enable this characterization of the metropolitan/non-metropolitan log wage gap:

$$\ln w_{Hm} - \ln w_{Ln} > \ln w_{Hm} - \ln w_{Hn} > \alpha \,. \qquad (4\text{-}7)$$

The wage gap for older workers exceeds the cost-of-living differential α because it reflects both the cost-of-living differential and the ability difference between metropolitan and non-metropolitan workers.

Cross-sectional estimation of a single metropolitan/non-metropolitan wage gap that pools the younger and older workers will tend to produce an estimate greater than α because it will average the α gap for younger workers with the greater-than-α gap for older workers. This accords with my finding in Section III that the cross-sectional wage gap exceeds the cost-of-living differential. Furthermore, an empirical analysis that allows the metropolitan/non-metropolitan wage gap to vary with experience will find that it increases with experience. Thus, my model explains that empirical regularity without resorting to Glaeser and Mare′′s claim that cities enable more rapid skill acquisition.

Instead, I explain the empirical pattern as the effect of dynamic ability sorting across sectors.

The longitudinal implications can be drawn from Table 9, which lists the sector and wage in each period for the high- and low-ability workers. Since my model abstracts from wage growth due to human capital accumulation, both the low-ability workers who stay in the non-metropolitan sector and the high-ability workers who stay in the metropolitan sector experience zero wage growth. Thus, my model, unlike Glaeser and Mare′'s, accords with my empirical finding in Section III of no difference in wage growth between metropolitan and non-metropolitan stayers.

High-ability workers who start in the non-metropolitan sector move to the metropolitan sector and experience a wage gain of $\ln w_{Hm} - \ln w_{Hn} > \alpha$, while low-ability workers who move from the metropolitan sector to the non-metropolitan sector experience a wage loss $\ln w_{Ln} - \ln w_{Lm}$ that is less than α in absolute value. Therefore, a simple longitudinal fixed-effects estimator that estimates a single metropolitan/non-metropolitan wage gap based on the pooled relative wage changes of the movers in both directions will estimate a weighted average of the above-α value for the high-ability movers and the below-α value based on the low-ability movers. This accords with my empirical finding in Section III that the fixed-effects estimate is considerably smaller than the cross-sectional estimate and approximately equals the cost-of-living differential. Furthermore, my model's combination of an above-α wage growth for non-metropolitan-to-metropolitan movers and a below-α wage loss for metropolitan-to-non-metropolitan movers accords remarkably well with the empirical pattern shown in the last column of Table 8.

V. Conclusion

In this paper, I have replicated and extended Glaeser and Mare′′s empirical analysis of the urban/non-urban wage gap, and I have taken issue with their interpretation of the evidence. My cross-sectional analysis indicates that the urban/non-urban wage gap is large, exceeds the cost-of-living differential, and is greater for more experienced workers. My longitudinal fixed-effects estimates that control for unobserved ability are smaller than the cross-sectional estimates and approximately equal the cost-of-living differential. Furthermore, my longitudinal evidence that urban stayers experience no faster wage growth than non-urban stayers contradicts Glaeser and Mare′′s theory that urban employment is conducive to more rapid skill acquisition.

My own theory accounts for the entire set of empirical regularities in terms of two highly plausible phenemena: (1) a compensating wage differential for the higher cost of living in cities and (2) a dynamic tendency for more able workers to gravitate to cities once they discover that they belong in the "big leagues."   Cross-sectional estimates exceed the cost-of-living differential because they also reflect a return to the greater ability of urban workers, and they increase with workers' experience because the sectoral sorting by ability occurs over time. Because the cross-sectional wage gap arises partly from the ability difference between urban and non-urban workers, longitudinal estimates that "difference out" the effects of unobserved ability come out smaller than cross-sectional estimates of the urban/non-urban wage gap.

References

Arellano, Manuel, “Computing Robust Standard Errors for Within-Groups Estimators,” Oxford Bulletin of Economics and Statistics, November 1987, 49, 431-434.

Citro, Constance F., and Robert T. Michael, eds., Measuring Poverty: A New Approach, Washington: National Academy Press, 1995.

Gibbons, Robert, and Lawrence Katz, “Does Unmeasured Ability Explain Inter-Industry Wage Differentials?” Review of Economic Studies, July 1992, 59, 515-535.

Glaeser, Edward L., and David C Mare′, “Cities and Skills,” Journal of Labor Economics, April 2001, 19, 316-342.

Griliches, Zvi, and Jerry A. Hausman, “Errors in Variables in Panel Data,” Journal of Econometrics, February 1986, 31, 93-118.

Harris, John R., and Michael P. Todaro, “Migration, Unemployment and Development: A Two-Sector Analysis,” American Economic Review, March 1970, 60, 126-142.

Johnson, D. Gale, “Comparability of Labor Capacities of Farm and Nonfarm Labor,” American Economic Review, June 1953, 43, 296-313.

Johnson, George, “Intermetropolitan Wage Differentials in the United States,” in The Measurement of Labor Cost, ed. Triplett, J.E., Chicago: University of Chicago Press, 1983.

Krugman, Paul, “Increasing Returns and Economic Geography,” Journal of Political Economy, June 1991, 99, 483-499.

Lucas, Jr., Robert E., “On the Mechanics of Economic Development,” Journal of Monetary Economics, July 1988, 22, 3-42.

Page, Marianne E., and Gary Solon, “Correlations between Brothers and Neighboring Boys in Their Adult Earnings: The Importance of Being Urban,” Journal of Labor Economics, October 2003, 21, 831-855.

Rauch, James, “Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities,” Journal of Urban Economics, November 1993, 34, 3-33.

Roback, Jennifer, “Wages, Rents, and the Quality of Life,” Journal of Political Economy, December 1982, 90, 1257-1278.

Rosen, Sherwin, “The Economics of Superstars,” American Economic Review, December 1981, 71, 845-858.

Table 1. Description of Variables

	Description of variables
PSID 1985-1991	
Log of Hourly Earnings	Log hourly wage rate of male head of family (V21484/V20344 in 1992 interview year in PSID)
Metropolitan	Metropolitan dummy equals to one if individual lives in the metropolitan area, otherwise zero (V20192 in 1991 interview year in PSID)[14]
North East	Regional dummy equals to one if individual lives in north-east region, otherwise zero (V20189 in 1991 interview year in PSID)
North Central	Regional dummy equals to one if individual lives in north-central region, otherwise zero (V20189 in 1991 interview year in PSID)
West	Regional dummy equals to one if individual lives in west region, otherwise zero (V20189 in 1991 interview year in PSID)
White	Race dummy equals to one if individual is white, otherwise zero (V20114 in 1991 interview year in PSID)
Experience	Potential experience (Age of head-years of completed education-6, maximum of age is 65) (Age variable is V30692 in 1991 interview year in PSID)
Education	Years of completed education (V30703 in 1991 interview year in PSID), Five education dummies: [0,9 years of schooling], [10, 11], [13,15: some college, but no degree], [16 years of schooling: Bachelor's degree], [more than 16 years of schooling: Master's degree, Professional degree, Doctorate degree]
1990 census 1percent IPUMS	
Log of Hourly Earnings	Log hourly wage rate of male head of household (INCOME1/WEEK89*HOUR89)
Metropolitan	Metropolitan dummy equals to one if individual lives in the metropolitan area, otherwise zero (MSA/PMSA codes 0040-9998 excluding 9997 of mixed areas, POWPUMA for the place of work)
North East	Regional dummy equals to one if individual lives in north-east region, otherwise zero (DIVISION: codes 1,2)
North Central	Regional dummy equals to one if individual lives in north-central region, otherwise zero (DIVISION: codes 3,4)
West	Regional dummy equals to one if individual lives in west region, otherwise zero (DIVISION: codes 8,9)
White	Race dummy equals to one if individual is white, otherwise zero (RACE)
Experience	Potential experience (Age of head-years of completed education-6, maximum of age is 65) (AGE-YEARSCH-6)
Education	Educational attainment (YEARSCH), Five education dummies: [0,9 years of schooling], [10, 11], [13,15: some college, but no degree], [16 years of schooling: Bachelor's degree], [more than 16 years of schooling: Master's degree, Professional degree, Doctorate degree]

[14] See section 3.1 for details of definitions of metropolitan areas.

Table 2. Sample Statistics

	Total (Means (S.D.))	Metro (Means (S.D.))	Non-metro (Means (S.D.))
1990 census (Place of residence)	N=337091	N=261708 (77.6%)	N=75383 (22.4%)
Log hourly wage	2.54(0.67)	2.61(0.66)	2.31(0.66)
White	0.873(0.33)	0.857(0.35)	0.927(0.26)
NE	0.19(0.39)	0.23((0.42)	0.05(0.22)
NC	0.26(0.44)	0.22(0.41)	0.43(0.49)
West	0.23(0.42)	0.24(0.43)	0.18(0.38)
Experience	23.38(11.17)	23.25(11.13)	23.85(11.30)
Education	13.18(2.75)	13.39(2.77)	12.49(2.57)
1990 census (Place of Work)	N=337091	N=268056 (79.5%)	N=69035 (20.5%)
Log hourly wage	2.54(0.67)	2.61(0.66)	2.30(0.67)
White	0.873(0.33)	0.859(0.34)	0.925(0.26)
NE	0.19(0.39)	0.22(0.42)	0.06(0.24)
NC	0.26(0.44)	0.22((0.41)	0.40(0.49)
West	0.23(0.42)	0.23(0.42)	0.18(0.38)
Experience	23.38(11.17)	23.29(11.13)	23.72(11.32)
Education	13.18(2.75)	13.34(2.77)	12.58(2.61)
1991 CPS (Place of residence)	N=20113	N=14748 (73.3%)	N=5365 (26.7%)
Log hourly wage	2.48(0.67)	2.56(0.64)	2.26(0.72)
White	0.89(0.32)	0.87(0.33)	0.92(0.27)
NE	0.21(0.41)	0.26(0.44)	0.08(0.28)
NC	0.27(0.44)	0.23(0.42)	0.37(0.48)
West	0.25(0.43)	0.26(0.44)	0.23(0.42)
Experience	20.82(11.20)	20.59(11.16)	21.45(11.28)
Education	13.34(2.97)	13.51(3.07)	12.86(2.64)
PSID 1985-1991 (Place of residence)	N=11585 (Person-year specific observations)	N=8260 (71.3%)	N=3325 (29.3%)
Log hourly wage	2.52(0.66)	2.64(0.64)	2.23(0.63)
White	0.93(0.40)	0.92(0.27)	0.95(0.21)
NE	0.20(0.46)	0.25(0.43)	0.09(0.28)
NC	0.31(0.37)	0.27(0.44)	0.40(0.49)
West	0.17(0.37)	0.20(0.40)	0.11(0.31)
Experience	19.34(9.96)	19.25(9.93)	19.60(10.02)
Education	13.59(2.38)	13.95(2.25)	12.73(2.48)

Table 3. Cross-Sectional Estimates of Metropolitan/Non-metropolitan Wage Gap

	1990 census(N=337091)		1991 March CPS(N=20113)		PSID (1985-1991)(N=1655)	
	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.32	Coef. (SE) R^2=0.34
Intercept	1.24 (0.013)[a]	1.24 (0.013)[a]	1.32 (0.034)[a]	1.33 (0.034)[a]	1.33 (0.149)[a]	1.35 (0.147)[a]
Metro	0.222 (0.0025)[a]		0.232 (0.010)[a]		0.308 (0.0292)[a]	
Large		0.270 (0.0027)[a]		0.252 (0.010)[a]		0.410 (0.0325)[a]
Small		0.161 (0.0029)[a]		0.190 (0.012)[a]		0.210 (0.0323)[a]
NE	0.138 (0.0030)[a]	0.132 (0.0030)[a]	0.118 (0.013)[a]	0.106 (0.013)[a]	0.107 (0.0356)[b]	0.083 (0.0354)[b]
NC	0.055 (0.0027)[a]	0.048 (0.0027)[a]	0.042 (0.012)[a]	0.036 (0.012)[a]	0.017 (0.0312)	-0.0063 (0.0309)
West	0.081 (0.0028)[a]	0.064 (0.0028)[a]	0.044 (0.012)[a]	0.037 (0.012)[a]	0.038 (0.0370)	0.0092 (0.0370)
White	0.147 (0.0031)[a]	0.159 (0.0031)[a]	0.125 (0.014)[a]	0.126 (0.014)[a]	0.210 (0.0477)[a]	0.225 (0.0477)[a]
Education [0,9]	-0.304 (0.0047)[a]	-0.306 (0.0047)[a]	-0.406 (0.018)[a]	-0.413 (0.018)[a]	-0.280 (0.0675)[a]	-0.291 (0.0666)[a]
Education [10,11]	-0.151 (0.0041)[a]	-0.152 (0.0041)[a]	-0.170 (0.019)[a]	-0.173 (0.019)[a]	-0.216 (0.0531)[a]	-0.206 (0.0524)[a]
Education [13,15]	0.125 (0.0027)[a]	0.121 (0.0027)[a]	0.161 (0.012)[a]	0.160 (0.012)[a]	0.132 (0.0329)[a]	0.118 (0.0325)[a]
Education [=16]	0.398 (0.0031)[a]	0.389 (0.0031)[a]	0.393 (0.014)[a]	0.390 (0.013)[a]	0.422 (0.0353)[a]	0.399 (0.0349)[a]
Education [>16]	0.555 (0.0037)[a]	0.544 (0.0037)[a]	0.478 (0.014)[a]	0.475 (0.014)[a]	0.526 (0.0380)[a]	0.503 (0.0376)[a]
Experience	0.080 (0.0025)[a]	0.079 (0.0025)[a]	0.084 (0.007)[a]	0.083 (0.007)[a]	0.069 (0.032)[b]	0.066 (0.031)[b]
Experience^2	-0.0026 (0.00017)[a]	-0.0025 (0.00017)[a]	-0.0034 (0.0005)[a]	-0.0033 (0.0005)[a]	-0.0027 (0.0023)	-0.0024 (0.0023)
Experience^3	0.000042 (0.0000046)[a]	0.000040 (0.0000046)[a]	0.000063 (0.000016)[a]	0.000061 (0.000016)[a]	0.000057 (0.000071)	0.000047 (0.000070)
Experience^4	-2.7E-7 (4.0E-8)[a]	-2.6E-7 (4.0E-8)[a]	-4.3E-7 (1.6E-7)[a]	-4.1E-7 (1.6E-7)[a]	-5.4E-7 (7.4E-7)	-4.2E-7 (7.3E-7)

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 4. Sensitivity of PSID Estimates to Sample Period and Estimation Method

	OLS	OLS with Huber-White Covariance	Between Estimation
PSID (1985-1991)	Coef.(SE) R^2=0.26, N=11585	Coef.(SE) R^2=0.26, N=11585	Coef.(SE) R^2=0.34
Time Dummies	Yes	Yes	No
Intercept	1.30(0.0023)[a]	1.30(0.149)[a]	1.33(0.149)[a]
Metro	0.295(0.012)[a]	0.295(0.027)[a]	0.308(0.029)[a]
NE	0.109(0.015)[a]	0.109(0.035)[b]	0.107(0.035)[b]
NC	0.016(0.013)	0.016(0.029)	0.017(0.031)
West	0.031(0.016)[c]	0.031(0.035)	0.038(0.037)
White	0.208(0.021)[a]	0.208(0.047)[a]	0.210(0.047)[a]
Education[0,9]	-0.290(0.029)[a]	-0.280(0.064)[a]	-0.280(0.067)[a]
Education[10,11]	-0.217(0.023)[a]	-0.217(0.049)[a]	-0.216(0.053)[a]
Education[13,15]	0.134(0.014)[a]	0.134(0.031)[a]	0.132(0.032)[a]
Education[=16]	0.424(0.015)[a]	0.424(0.036)[a]	0.422(0.035)[a]
Education[>16]	0.528(0.016)[a]	0.528(0.041)[a]	0.526(0.038)[a]
Experience	0.050(0.0059)[a]	0.056(0.0074)[a]	0.069(0.032)[b]
$Experience^2$	-0.0011(0.00039)[a]	-0.0011(0.00038)[a]	-0.0027(0.0023)
$Experience^3$	6.9E-6(0.000012)	6.9E-6(0.000014)	0.000057(0.000071)
$Experience^4$	-6.0E-9(0.000012)	-6.0E-9(1.97E-7)	-5.4E-7(7.4E-7)
PSID (1968-1985)	Coef.(SE) R^2=0.40, N=35685	Coef.(SE) R^2=0.40, N=35685	Coef.(SE) R^2=0.36, N=3998
Time Dummies	Yes	Yes	No
Intercept	0.505(0.020)[a]	0.505(0.020)[a]	1.14(0.040) [a]
Metro	0.242(0.006)[a]	0.242(0.015)[a]	0.294(0.021)[a]
NE	0.129(0.009)[a]	0.129(0.021)[a]	0.206(0.027)[a]
NC	0.086(0.007)[a]	0.086(0.019)[a]	0.140(0.023)[a]
West	0.076(0.009)[a]	0.076(0.022)[a]	0.136(0.028)[a]
White	0.150(0.011)[a]	0.150(0.024)[a]	0.408(0.028)[a]
Education[0,9]	-0.257(0.010)[a]	-0.257(0.027)[a]	-0.108(0.026)[a]
Education[10,11]	-0.120(0.011)[a]	-0.120(0.026)[a]	-0.100(0.040)[a]
Education[13,15]	0.117(0.009)[a]	0.117(0.022)[a]	0.092(0.033)[a]
Education[=16]	0.344(0.008)[a]	0.344(0.021)[a]	0.243(0.035)[a]
Education[>16]	0.452(0.022)[a]	0.452(0.040)[a]	1.439(0.236)[a]
Experience	0.028(0.007)[a]	0.028(0.0034)[a]	0.008(0.0021)[a]
$Experience^2$	-0.0002(0.000019)[a]	-0.0002(0.000092)[b]	-0.00018(0.000047)[a]
$Experience^3$	-0.0000062(0.0000001)[a]	-0.0000062(4.3E-7)[a]	0.000003(0.0000004)[a]
$Experience^4$	2.57E-8(4.0E-10)[a]	2.57E-8(1.71E-8)	1.9E-8(1.0E-8)[b]

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 5. Cross-Sectional Estimates with Adjustments for Cost of Living

	1990 census		1991 CPS		PSID (1985-1991)	
	LHS	RHS	LHS	RHS	LHS	RHS
	Coef. (SE) R^2=0.19	Coef. (SE) R^2=0.22	Coef. (SE) R^2=0.18	Coef. (SE) R^2=0.22	Coef. (SE) R^2=0.27	Coef. (SE) R^2=0.33
Intercept	1.33 (0.013)[a]	1.39 (0.013)[a]	1.42 (0.034)[a]	1.46 (0.036)[a]	1.44 (0.148)[a]	1.68 (0.15)[a]
Metro	0.129 (0.0025)[a]	0.070 (0.0036)[a]	0.126 (0.0099)[a]	0.092 (0.017)[a]	0.202 (0.028)[a]	-0.011 (0.056)
NE	0.087 (0.0030)[a]	0.055 (0.0033)[a]	0.047 (0.013)[a]	0.024 (0.016)	0.029 (0.035)	-0.126 (0.050)[b]
NC	0.049 (0.0027)[a]	0.046 (0.0027)[a]	0.037 (0.012)[a]	0.036 (0.012)[a]	0.00062 (0.030)	-0.030 (0.031)
West	0.033 (0.0028)[a]	0.0036 (0.0031)	-0.002 (0.012)	-0.017 (0.012)	-0.014 (0.037)	-0.121 (0.044)
White	0.159 (0.0031)[a]	0.166 (0.0031)[a]	0.131 (0.014)[a]	0.133 (0.014)[a]	0.215 (0.047)[a]	0.223 (0.047)[a]
Education[0,9]	-0.306 (0.0047)[a]	-0.307 (0.0047)[a]	-0.416 (0.018)[a]	-0.419 (0.018)[a]	-0.280 (0.067)[a]	-0.280 (0.067)[a]
Education[10,11]	-0.153 (0.0041)[a]	-0.154 (0.0041)[a]	-0.173 (0.019)[a]	-0.174 (0.019)[a]	-0.214 (0.052)[a]	-0.209 (0.052)[a]
Education[13,15]	0.122 (0.0027)[a]	0.119 (0.0027)[a]	0.159 (0.012)[a]	0.159 (0.012)[a]	0.129 (0.032)[a]	0.118 (0.032)[a]
Education[=16]	0.391 (0.0031)[a]	0.387 (0.0031)[a]	0.390 (0.014)[a]	0.390 (0.014)[a]	0.415 (0.035)[a]	0.398 (0.035)[a]
Education[>16]	0.546 (0.0037)[a]	0.541 (0.0037)[a]	0.473 (0.014)[a]	0.471 (0.014)[a]	0.518 (0.037)[a]	0.502 (0.037)[a]
Experience	0.079 (0.0025)[a]	0.079 (0.0025)[a]	0.084 (0.007)[a]	0.084 (0.007)[a]	0.067 (0.031)[b]	0.065 (0.031)[b]
$Experience^2$	-0.0026 (0.00017)[a]	-0.0025 (0.00017)[a]	-0.0034 (0.0005)[a]	-0.0034 (0.0005)[a]	-0.0027 (0.0023)	-0.0024 (0.0023)
$Experience^3$	0.000040 (4.6E-6)[a]	0.000040 (4.6E-6)[a]	0.000062 (0.000016)[a]	0.000062 (0.000016)[a]	0.000053 (0.000071)	0.000045 (0.000071)
$Experience^4$	-2.6E-7 (4.0E-8)[a]	-2.6E-7 (4.0E-8)[a]	-4.2E-7 (1.6E-7)[a]	-4.2E-7 (1.6E-7)[a]	-4.8E-7 (7.3E-7)	-3.9E-7 (7.3E-7)
Cost of living		1.237 (0.021)[a]		1.287 (0.128)[a]		3.019 (0.461)[a]

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 6. Cross-Sectional Estimates with Finer Categorizations of Metropolitan Status

	1990 census (N=337091)		1991CPS (N=20113)	PSID (1985-1991) (N=1655)
	Place of residence	Place of work	Place of residence	Place of residence
	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.21	Coef. (SE) R^2=0.34
Intercept	1.26 (0.013)[a]	1.24 (0.013)[a]	1.36 (0.034)[a]	1.36 (0.147)[a]
Central City	0.178 (0.0036)[a]	0.258 (0.0031)[a]	0.154 (0.012)[a]	0.377 (0.0379)[a]
Suburban	0.253 (0.0026)[a]	0.228 (0.0027)[a]	0.277 (0.010)[a]	0.449 (0.0395)[a]
Other Metropolitan Areas	0.103 (0.0042)[a]	0.093 (0.0042)[a]		0.211 (0.0322)[a]
NE	0.130 (0.0030)[a]	0.136 (0.0030)[a]	0.114 (0.013)[a]	0.079 (0.0354)[b]
NC	0.052 (0.0027)[a]	0.052 (0.0027)[a]	0.039 (0.012)[a]	-0.0053 (0.0309)
West	0.078 (0.0028)[a]	0.077 (0.0028)[a]	0.048 (0.012)[a]	0.015 (0.0370)
White	0.141 (0.0032)[a]	0.153 (0.0031)[a]	0.098 (0.014)[a]	0.214 (0.0477)[a]
Education[0,9]	-0.298 (0.0047)[a]	-0.303 (0.0047)[a]	-0.389 (0.018)[a]	-0.294 (0.0666)[a]
Education[10,11]	-0.149 (0.0041)[a]	-0.150 (0.0041)[a]	-0.166 (0.019)[a]	-0.206 (0.0524)[a]
Education[13,15]	0.122 (0.0027)[a]	0.124 (0.0027)[a]	0.158 (0.012)[a]	0.121 (0.0325)[a]
Education[=16]	0.393 (0.0031)[a]	0.395 (0.0031)[a]	0.391 (0.013)[a]	0.401 (0.0349)[a]
Education[>16]	0.550 (0.0037)[a]	0.552 (0.0037)[a]	0.477 (0.014)[a]	0.505 (0.0376)[a]
Experience	0.078 (0.0025)[a]	0.078 (0.0025)[a]	0.082 (0.007)[a]	0.066 (0.031)[b]
Experience^2	-0.0025 (0.00017)[a]	-0.0025 (0.00017)[a]	-0.0033 (0.0005)[a]	-0.0025 (0.0023)
Experience^3	0.000039 (4.6E-6)[a]	0.000039 (4.6E-6)[a]	0.000061 (0.000016)[a]	0.000050 (0.000070)
Experience^4	-2.6E-7 (4.0E-8)[a]	-2.6E-7 (4.0E-8)[a]	-4.1E-7 (1.6E-7)[a]	-4.4E-7 (7.3E-7)

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 7. Interaction of Urban/Non-urban Wage Gap with Education and Experience

	1990 census (N=337091)		1991CPS (N=20113)	PSID (1985-1991) (N=1655)
	Place of Residence	Place of Work	Place of Residence	Place of Residence
	Coef.(SE), R^2=0.21	Coef(SE), R^2=0.21	Coef(SE), R^2=0.21	Coef(SE),R^2=0.32
Intercept	1.40	1.40	1.45	1.36
	(0.027)[a]	(0.027)[a]	(0.063)[a]	(0.273)[a]
D (Metro)	0.012	0.015	0.056	0.234
	(0.030)	(0.031)	(0.072)	(0.324)
NE	0.139	0.142	0.120	0.108
	(0.0030)[a]	(0.0030)[a]	(0.012)[a]	(0.035)[a]
NC	0.056	0.051	0.045	0.015
	(0.0027)[a]	(0.0027)[a]	(0.011)[a]	(0.031)
West	0.084	0.084	0.049	0.038
	(0.0028)[a]	(0.0028)[a]	(0.011)[a]	(0.037)
White	0.145	0.144	0.126	0.208
	(0.0031)[a]	(0.0031)[a]	(0.013)[a]	(0.048)[a]
Edu[0,9]	-0.234	-0.236	-0.331	-0.262
	(0.0089)[a]	(0.0093)[a]	(0.035)[a]	(0.088)[a]
Edu[10,11]	-0.143	-0.140	-0.131	-0.143
	(0.0077)[a]	(0.0081)[a]	(0.031)[a]	(0.086)[c]
Edu[13,15]	0.082	0.092	0.113	0.103
	(0.0055)[a]	(0.0058)[a]	(0.022)[a]	(0.066)
Edu[=16]	0.274	0.305	0.283	0.244
	(0.0074)[a]	(0.0075)[a]	(0.027)[a]	(0.073)[a]
Edu[>16]	0.399	0.427	0.355	0.465
	(0.0093)[a]	(0.0094)[a]	(0.035)[a]	(0.085)[a]
Exp	0.064	0.0623	0.064	0.090
	(0.0025)[a]	(0.0055)[a]	(0.014)[a]	(0.061)
Exp^2	-0.0019	-0.0018	-0.0020	-0.0050
	(0.00017)[a]	(0.00037)[a]	(0.0011)[c]	(0.0044)
Exp^3	0.000029	0.000024	0.000025	0.00013
	(0.0000046)[a]	(0.000010)[b]	(0.000033)	(0.00013)
Exp^4	-2.0E-7	-1.50E-7	-1.2E-7	-1.3E-6
	(9.0E-8)[b]	(9.0E-8)[c]	(3.3E-7)	(1.3E-6)
D*Edu[0,9]	-0.0939	-0.088	-0.097	-0.053
	(0.0104)[a]	(0.0108)[a]	(0.041)[b]	(0.146)
D*Edu[10,11]	-0.0093	-0.014	-0.059	-0.124
	(0.0091)	(0.0094)	(0.039)	(0.111)
D*Edu[13,15]	0.0584	0.046	0.067	0.048
	(0.0063)[a]	(0.0066)[a]	(0.026)[a]	(0.078)
D*Edu[=16]	0.1538	0.1211	0.142	0.238
	(0.0082)[a]	(0.0083)[a]	(0.031)[a]	(0.085)[a]
D*Edu[>16]	0.1889	0.1598	0.158	0.088
	(0.0102)[a]	(0.0102)[a]	(0.034)[a]	(0.097)
D*Exp	0.0199	0.0225	0.027	-0.026
	(0.0061)[a]	(0.0062)[a]	(0.016)	(0.073)
D*Exp^2	-0.00083	-0.0010	-0.0018	0.0032
	(0.00041)[b]	(0.00042)[b]	(0.0012)	(0.0054)
D*Exp^3	0.0000146	0.0000216	0.000049	-0.00011
	(0.000011)	(0.000011)[c]	(0.000038)	(0.00015)
D*Exp^4	-7.8E-8	-1.5E-7	-4.1E-7	1.3E-6
	(1.0E-7)	(1.0E-7)	(3.8E-7)	(1.6E-6)

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 8. Longitudinal Estimates from the PSID

	Covariance Estimation, N=11585		Long Difference Estimation (1985-1991), (N=1655)	
		With Controlling Geographic Price Differentials		
	Coef(SE) R^2=0.10	Coef(SE) R^2=0.10	Coef(SE) R^2=0.044	Coef(SE) R^2=0.045
Intercept	-6.44 (0.246)[a]	-6.43 (0.246)[a]	0.63 (0.097)[a]	0.63 (0.101)[a]
Time Dummies	Yes	Yes	No	No
D (Metro)	0.113 (0.024)[a]	0.008 (0.024)	0.142 (0.061)[b]	
(D_{nm})				0.187 (0.091)[b]
(D_{mn})				-0.097 (0.091)
(D_{mm})				-0.013 (0.033)
NE	0.033 (0.046)	-0.042 (0.047)	-0.030 (0.118)	-0.030 (0.118)
NC	0.017 (0.032)	-0.010 (0.033)	0.092 (0.084)	0.092 (0.084)
West	-0.091 (0.037)	-0.172 (0.039)	-0.134 (0.089)	-0.134 (0.089)
Experience^2	-6.7E-7 (6.1E-6)	-6.3E-7 (6.1E-6)	-0.0023 (0.0013)[c]	-0.0023 (0.0013)[c]
Experience^3	-0.000038 (0.0000051)[a]	-0.000038 (0.0000051)[a]	0.000027 (0.000042)	0.000027 (0.000042)
Experience^4	5.2E-7 (9.0E-9)[a]	5.2E-7 (9.0E-9)[a]	-1.2E-7 (4.5E-7)	-1.2E-7 (4.5E-7)

Superscripts a, b, and c represent statistical significance at, respectively, the 1, 5, and 10 percent levels.

Table 9. Summary of Model's Implications for Log Wages in Each Period

Location in 2nd period	Change of location	Ability Type	1st period log wage	2nd period log wage	Δ log wage
Metro	Stayer	High	$\ln w_{Hm}$	$\ln w_{Hm}$	0
	Mover	High	$\ln w_{Hn}$	$\ln w_{Hm}$	$\ln w_{Hm} - \ln w_{Hn} > \alpha$
Non-metro	Stayer	Low	$\ln w_{Ln}$	$\ln w_{Ln}$	0
	Mover	Low	$\ln w_{Lm}$	$\ln w_{Ln}$	$\ln w_{Ln} - \ln w_{Lm} > -\alpha$

Figure 1.

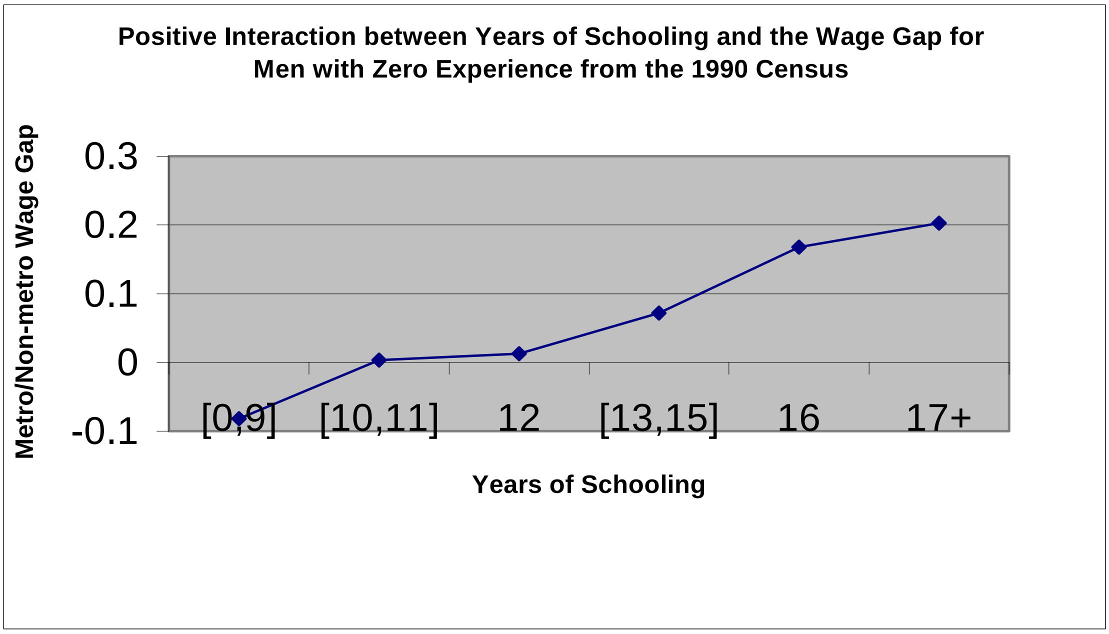
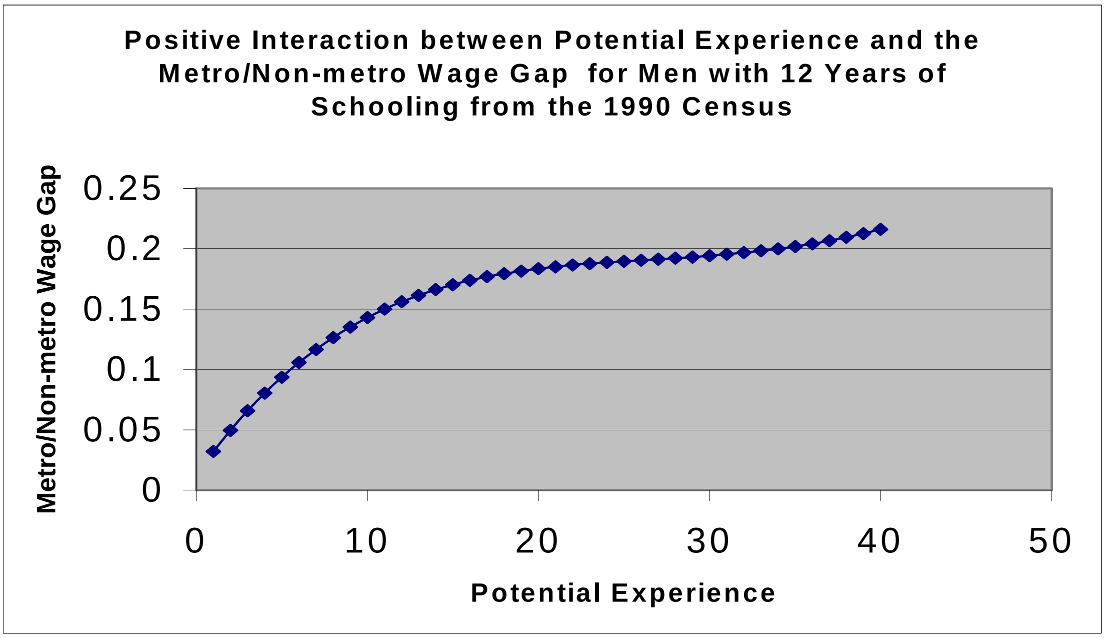

Positive Interaction between Years of Schooling and the Wage Gap for Men with Zero Experience from the 1990 Census
Metro/Non-metro Wage Gap
0.3
0.2
0.1
0
-0.1
[0,9]
[10,11]
12
[13,15]
16
17+
Years of Schooling

Figure 2.



Positive Interaction between Potential Experience and the Metro/Non-metro Wage Gap for Men with 12 Years of Schooling from the 1990 Census
Metro/Non-metro Wage Gap
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
Potential Experience